

ON A SPECTRAL ESTIMATE OBTAINED BY AN
AUTOREGRESSIVE MODEL FITTING

BY

MITUAKI HUZII

TECHNICAL REPORT NO. 22
FEBRUARY 1976

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA



ON A SPECTRAL ESTIMATE OBTAINED BY AN
AUTOREGRESSIVE MODEL FITTING

BY

MITUAKI HUZII

TECHNICAL REPORT NO. 22

FEBRUARY 1976

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)

OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

On a Spectral Estimate Obtained by an
Autoregressive Model Fitting

by

Mituaki Huzii

Tokyo Institute of Technology and
Stanford University

1. Introduction

Let $X(t)$ be a real-valued stationary normal process with a discrete time parameter t . For simplicity, we shall assume $EX(t) = 0$ and denote $R(h) = EX(t+h)X(t)$ and $\rho_h = R(h)/R(0)$. Let us assume $X(t)$ is observed at $t = 1, 2, \dots, N$ and has a spectral density $f(\lambda)$, where λ is a frequency parameter and $-1/2 \leq \lambda \leq 1/2$.

In this paper, we shall discuss efficient estimation of $f(\lambda)$. If we can assume $X(t)$ is an autoregressive process of order K , K being a known positive integer, we can obtain, easily, the maximum likelihood estimate of $f(\lambda)$ when N is sufficiently large. But for an actual process, we usually do not know the value of K . Recently, the estimates obtained by fitting an autoregressive model have been developed and discussed by many authors, i.e., Akaike [1], Parzen [8], Gersch and Sharpe [4] and Jones [7].

In this paper, we treat a process expressed as an autoregressive process of infinite order satisfying some conditions. We construct

an estimate by fitting an autoregressive model of finite order K . In Section 2 we discuss the asymptotic bias of this estimate for a fixed K when N tends to infinity. In Section 3 we consider K as a function of N , tending to infinity as N tends to infinity. Berk [3] has discussed a similar situation. He has shown the consistency and the asymptotic normality of the estimate when N tends to infinity. Although we shall discuss the statistical properties of the same estimate, the process under consideration here satisfies stronger conditions than his, and under our conditions we show that this estimate has a property of efficiency as N tends to infinity.

2. An Autoregressive Approximation and Bias.

In the following we shall assume, furthermore, $X(t)$ satisfies the following assumption:

Assumption 1. $X(t)$ satisfies the relation

$$(1) \quad \sum_{k=0}^{\infty} a_k X(t-k) = \xi(t), \quad a_0 = 1,$$

where $\{\xi(t)\}$ are mutually independent random variables each of which
has the distribution $N(0, \sigma_{\xi}^2)$ and $\{a_k\}$ are constants such that

$$|a_k| \leq \alpha^k, \quad 0 < \alpha < 1/2,$$

for every $k \geq 1$.

In this case, $X(t)$ has a backward moving average representation

$$X(t) = \sum_{k=0}^{\infty} G_k \xi(t-k),$$

where $\{G_k\}$ are constants.

As is shown in Huzii [6], we have the following result under the above assumptions.

Lemma 1. We have

$$|\rho_h| \leq C(2\alpha)^h \quad \text{and} \quad |G_h| \leq (2\alpha)^{h/2}.$$

for any h , $h \geq 1$, where $C = 1/2^2(1-(2\alpha)^2)$.

Now we shall discuss the estimation of the spectral density $f(\lambda)$. We shall regard $X(t)$ as an autoregressive process of order K , K being a positive integer, and obtain $\{a_k^{(K)}; 1 \leq k \leq K\}$ which minimize

$$Q = \sum_{t=K+1}^N \left(X(t) + \sum_{k=1}^K a_k^{(K)} X(t-k) \right)^2.$$

Let us denote $a_k^{(K)}$, which minimizes Q , as $\hat{a}_k^{(K)}$ for $1 \leq k \leq K$. If $X(t)$ is an autoregressive process of order K , the $\hat{a}_k^{(K)}$'s are asymptotically maximum likelihood estimates of autoregressive coefficients. But here $X(t)$ is not an autoregressive process of finite order. So this argument does not hold. The $a_k^{(K)}$'s are the solutions of the simultaneous equation

$$(2) \quad \sum_{k=1}^K \hat{a}_k^{(K)} \hat{R}(k, \ell) = -\hat{R}(0, \ell), \quad 1 \leq \ell \leq K,$$

where

$$\hat{R}(k, \ell) = \frac{1}{N-K} \sum_{t=K+1}^N X(t-k) X(t-\ell).$$

Using these $\hat{a}_k^{(K)}$'s, we shall construct the estimate $\hat{f}(\lambda)$ of $f(\lambda)$ as follows

$$\hat{f}^{(K)}(\lambda) = \frac{\hat{\sigma}_{\xi}^2(K)}{\left(\sum_{k=0}^K \hat{a}_k^{(K)} \cos 2\pi k \lambda \right)^2 + \left(\sum_{k=1}^K \hat{a}_k^{(K)} \sin 2\pi k \lambda \right)^2},$$

where $\hat{a}_0^{(K)} = 1$ and

$$\hat{\sigma}_{\xi}^2(K) = \frac{1}{N-K} \sum_{t=K+1}^N \left(X(t) + \sum_{k=1}^K \hat{a}_k^{(K)} X(t-k) \right)^2.$$

Now let us evaluate the bias of the estimate $\hat{f}^{(K)}(\lambda)$ for a sufficiently large and fixed K . We shall denote $\hat{a}^{(K)} = (\hat{a}_1^{(K)}, \hat{a}_2^{(K)}, \dots, \hat{a}_K^{(K)})'$ and $\hat{R}_K = (\hat{R}(0,1), \hat{R}(0,2), \dots, \hat{R}(0,K))'$. And let $\hat{\tilde{R}}$ be the $K \times K$ matrix whose (k, ℓ) element is $\hat{R}(k, \ell)$.

Then the simultaneous equation (2) can be written

$$(3) \quad \hat{\tilde{R}} \hat{a}^{(K)} = -\hat{R}_K.$$

Now let Q_J be the $J \times J$ matrix whose (k, ℓ) element is $\rho_{k-\ell}$. Using the result of Lemma 1, we can show $\hat{R}(k, \ell)$ converges in probability to $R(k - \ell)$ as N tends to infinity. Let us denote the $K \times K$ matrix, whose (k, ℓ) element is $R(k - \ell)$, as \tilde{R} . Then for a sufficiently large and fixed K , every element of $\hat{\tilde{R}}$ converges in probability to the corresponding element of \tilde{R} . Let us put $\tilde{R}_K = (R(1), R(2), \dots, R(K))'$. Then every element of $\hat{\tilde{R}}_K$ converges

in probability to the corresponding element of \tilde{R}_K . We can consider $|\tilde{R}| \neq 0$ for any K , $K \geq 1$. Using the relation (3), we can show the distribution of $\sqrt{N} (\hat{a}(K) + \tilde{R}^{-1} \tilde{R}_K)$ converges to the normal distribution with mean vector 0 and a finite covariance matrix. This can be shown by the same method as in Anderson [2], Chap. 5. Using this result, we shall evaluate the asymptotic bias of the estimate $\hat{f}^{(K)}(\lambda)$ as $N \rightarrow \infty$. In the first place, we shall evaluate the value $\tilde{a}(K) = -\tilde{R}^{-1} \tilde{R}_K$. Let us put $\tilde{a}(K) = (a_1^{(K)}, a_2^{(K)}, \dots, a_K^{(K)})'$. Then we have the following lemma.

Lemma 2. We have, for $1 \leq k \leq K$,

$$|a_k^{(K)} - a_k| \leq C_1 \alpha^K,$$

where C_1 is a constant, being independent of K .

Proof. Let us put $\rho_K = (\rho_1, \rho_2, \dots, \rho_K)'$ for simplicity. Then we have $\tilde{a}(K) = -\tilde{Q}_K^{-1} \rho_K$. We can express, by using the result of Wise [8], the (k, ℓ) element $q_{k,\ell}^{-1}$ of \tilde{Q}_K^{-1} as $\tilde{q}_{k,\ell}^{-1} + \varepsilon_{k,\ell}^{(K)}$, where

$$\tilde{q}_{k,\ell}^{-1} = \begin{cases} \sum_{j=1}^{\ell} a_{k-j} a_{\ell-j}, & \text{if } k > \ell \\ 1 + \sum_{j=1}^{k-1} a_j^2, & \text{if } k = \ell \\ \tilde{q}_{\ell,k}^{-1}, & \text{if } k < \ell \end{cases}$$

and $\varepsilon_{k,\ell}^{(K)}$ is uniformly bounded for either $k < K - K_0$ or

$l < K - K_0$, K_0 being a fixed positive integer, when K is sufficiently large. So taking into account of Assumption 1, we have

$$|q_{k,l}^{-1}| \leq q,$$

where q is a constant and independent of K . Now we have the relation

$$\sum_{k=0}^{\infty} a_k \rho_{k-j} = 0, \quad j = 1, 2, 3, \dots,$$

by (1). This can be written as

$$(4) \quad \sum_{k=1}^K a_k \rho_{k-j} = -\rho_j - \sum_{k=K+1}^{\infty} a_k \rho_{k-j}, \quad j = 1, 2, \dots$$

Putting $\tilde{a} = (a_1, a_2, \dots, a_K)'$, $\delta_j = \sum_{k=K+1}^{\infty} a_k \rho_{k-j}$ and $\tilde{\Delta} = (\delta_1, \delta_2, \dots, \delta_K)'$, we have

$$(5) \quad \tilde{a} = -Q_K^{-1} \tilde{\rho}_K - Q_K^{-1} \tilde{\Delta}$$

by using (4). From this relation, we obtain

$$\tilde{a}(K) - \tilde{a} = Q_K^{-1} \tilde{\Delta}.$$

Let us evaluate δ_j for $1 \leq j \leq K$.

$$\begin{aligned}
|\delta_j| &\leq \sum_{k=K+1}^{\infty} |a_k| |\rho_{k-j}| \\
&\leq \sum_{k=K+1}^{\infty} \alpha^k C (2\alpha)^{k-j} = C \frac{\alpha^{K+1} (2\alpha)^{K+1-j}}{1 - 2\alpha^2} .
\end{aligned}$$

So we have

$$\begin{aligned}
|a_k^{(K)} - a_k| &\leq C q \sum_{j=1}^K \frac{\alpha^{K+1} (2\alpha)^{K+1-j}}{1 - 2\alpha^2} \\
&\leq \frac{q C \alpha^{K+1} (2\alpha)}{(1 - 2\alpha^2)(1 - 2\alpha)} = C_1 \alpha^K ,
\end{aligned}$$

for $1 \leq k \leq K$, where $C_1 = 2qC\alpha^2/(1 - 2\alpha^2)(1 - 2\alpha)$. We can easily find that, when N tends to infinity, the mean value of the limiting distribution of $\sqrt{N}(\hat{f}^{(K)}(\lambda) - f^{(K)}(\lambda))$ is zero, where

$$f^{(K)}(\lambda) = \frac{\sigma_{\xi}^2(K)}{\left(\sum_{k=0}^K a_k^{(K)} \cos 2\pi k \lambda \right)^2 + \left(\sum_{k=1}^K a_k^{(K)} \sin 2\pi k \lambda \right)^2}$$

and

$$\sigma_{\xi}^2(K) = R(0) + \sum_{k=1}^K a_k^{(K)} R(k) .$$

Now we have

$$\sigma_{\xi}^2 - \sigma_{\xi}^2(K) = \sum_{k=K+1}^{\infty} a_k R(k) - \sum_{k=1}^K (a_k^{(K)} - a_k) R(k)$$

and

$$\begin{aligned}
& \left(\sum_{k=0}^{\infty} a_k \cos 2\pi k\lambda \right)^2 + \left(\sum_{k=1}^{\infty} a_k \sin 2\pi k\lambda \right)^2 \\
& - \left(\sum_{k=0}^K a_k^{(K)} \cos 2\pi k\lambda \right)^2 - \left(\sum_{k=1}^K a_k^{(K)} \sin 2\pi k\lambda \right)^2 \\
& = \left(\sum_{k=K+1}^{\infty} a_k \cos 2\pi k\lambda \right)^2 + \left(\sum_{k=K+1}^{\infty} a_k \sin 2\pi k\lambda \right)^2 \\
& + 2 \left(\sum_{k=0}^K a_k \cos 2\pi k\lambda \right) \left(\sum_{k=K+1}^{\infty} a_k \cos 2\pi k\lambda \right) \\
& + 2 \left(\sum_{k=1}^K a_k \sin 2\pi k\lambda \right) \left(\sum_{k=K+1}^{\infty} a_k \sin 2\pi k\lambda \right) \\
& - 2 \left(\sum_{k=0}^K a_k \cos 2\pi k\lambda \right) \left(\sum_{k=0}^K (a_k^{(K)} - a_k) \cos 2\pi k\lambda \right) \\
& - 2 \left(\sum_{k=1}^K a_k \sin 2\pi k\lambda \right) \left(\sum_{k=1}^K (a_k^{(K)} - a_k) \sin 2\pi k\lambda \right) \\
& - \left(\sum_{k=0}^K (a_k^{(K)} - a_k) \cos 2\pi k\lambda \right)^2 \\
& - \left(\sum_{k=1}^K (a_k^{(K)} - a_k) \sin 2\pi k\lambda \right)^2 .
\end{aligned}$$

And, also, we have, for example,

$$\left| \sum_{k=K+1}^{\infty} a_k \cos 2\pi k\lambda \right| \leq \alpha^{K+1} / (1 - \alpha) .$$

and

$$\left| \sum_{k=0}^K (a_k^{(K)} - a_k) \cos 2\pi k \lambda \right| \leq C_1 K \alpha^K.$$

Therefore, we have

$$\begin{aligned} & \left| \left\{ \left(\sum_{k=0}^K a_k^{(K)} \cos 2\pi k \lambda \right)^2 + \left(\sum_{k=0}^K a_k^{(K)} \sin 2\pi k \lambda \right)^2 \right\} \right. \\ & \quad \left. - \left\{ \left(\sum_{k=0}^{\infty} a_k \cos 2\pi k \lambda \right)^2 + \left(\sum_{k=0}^{\infty} a_k \sin 2\pi k \lambda \right)^2 \right\} \right| \\ & \leq C_2 K \alpha^K, \end{aligned}$$

where C_2 is a constant being independent of K . So we can show

$$|f^{(K)}(\lambda) - f(\lambda)| \leq C_3 K \alpha^K,$$

where C_3 is a constant being independent of K . We obtain the following theorem.

Theorem 1. Let $X(t)$ be a stationary normal process satisfying the Assumption 1. Then

$$|f^{(K)}(\lambda) - f(\lambda)| \leq C_3 K \alpha^K$$

for sufficiently large K , where C_3 is a constant, independent
of K .

3. Asymptotic Efficiency of $\hat{f}^{(K)}(\lambda)$ in a Sense.

In this section, we shall consider K to be a function of N . Recently, Berk [3] has shown that $\hat{f}^{(K)}(\lambda)$ is consistent and asymptotically normal when N tends to infinity under the condition that K tends to infinity and K^3/N tends to 0 and some other conditions. Here we make the following assumption.

Assumption 2. K is a function of N such that K tends to infinity and K^4/N and $N^9 K^2 \alpha^K$ tend to zero when N tends to infinity.

This assumption is stronger than Berk's condition. We shall define $f_K(\lambda)$ as

$$f_K(\lambda) = \frac{\sigma_\xi^2}{\left(\sum_{k=0}^K a_k \cos 2\pi k \lambda \right)^2 + \left(\sum_{k=1}^K a_k \sin 2\pi k \lambda \right)^2}.$$

Then we have

$$|f_K(\lambda) - f(\lambda)| \leq C_4 \alpha^K,$$

where C_4 is a constant being independent of K . In the following we shall consider the estimation of $f_K(\lambda)$ instead of $f(\lambda)$. We have

$$|f^{(K)}(\lambda) - f_K(\lambda)| \leq C_2' K \alpha^K,$$

where C_2' is a constant, independent of K . Now we have

$$\begin{aligned} E\left(\hat{f}^{(K)}(\lambda) - f(\lambda)\right)^2 &= E\left(\hat{f}^{(K)}(\lambda) - f^{(K)}(\lambda)\right)^2 \\ &\quad + \left(f^{(K)}(\lambda) - f_K(\lambda)\right)^2 + \left(f_K(\lambda) - f(\lambda)\right)^2 \\ &\quad + 2\left(f^{(K)}(\lambda) - f_K(\lambda)\right) E\left(\hat{f}^{(K)}(\lambda) - f^{(K)}(\lambda)\right) \\ &\quad + 2\left(f^{(K)}(\lambda) - f_K(\lambda)\right)\left(f_K(\lambda) - f(\lambda)\right) \\ &\quad + 2\left(f_K(\lambda) - f(\lambda)\right) E\left(\hat{f}^{(K)}(\lambda) - f^{(K)}(\lambda)\right). \end{aligned}$$

But when we consider the case in which N is sufficiently large, we can ignore the terms

$$2\left(f^{(K)}(\lambda) - f_K(\lambda)\right) E\left(\hat{f}^{(K)}(\lambda) - f^{(K)}(\lambda)\right)$$

and

$$2\left(f_K(\lambda) - f(\lambda)\right) E\left(\hat{f}^{(K)}(\lambda) - f^{(K)}(\lambda)\right)$$

by comparing with the other terms. And we know

$$\begin{aligned} &\left| \left(f^{(K)}(\lambda) - f_K(\lambda)\right)^2 + \left(f_K(\lambda) - f(\lambda)\right)^2 \right. \\ &\quad \left. + 2\left(f^{(K)}(\lambda) - f_K(\lambda)\right)\left(f_K(\lambda) - f(\lambda)\right) \right| \leq C_5(K \alpha^K)^2, \end{aligned}$$

where C_5 is a constant, independent of K . When we consider

$$\lim_{N \rightarrow \infty} \frac{N}{K} E \left(\hat{f}^{(K)}(\lambda) - f(\lambda) \right)^2,$$

we have

$$\lim_{N \rightarrow \infty} \frac{N}{K} C_5 (K \alpha^K)^2 = 0$$

by Assumption 2. Using the result of Berk [3], we can see that the distribution function of $\sqrt{N/K} \left(\hat{f}^{(K)}(\lambda) - f^{(K)}(\lambda) \right)$ (or $\sqrt{N/K} \left(\hat{f}^{(K)}(\lambda) - f(\lambda) \right)$) tends to the normal distribution function with mean 0 and variance $2f(\lambda)^2$ or $4f(\lambda)^2$ corresponding to when $0 < \lambda < 1/2$ or when $\lambda = 0$ or $1/2$, respectively.

Now let us define an autoregressive process of order K . Let $X_K(t)$ be a stationary normal process with mean 0 and satisfy the relation

$$\sum_{k=0}^K a_k X_K(t-k) = \xi(t),$$

where $\{a_k\}$ and $\xi(t)$ are the same as those in (1) and $a_0 = 1$.

We shall put $R_K(h) = EX_K(t+h)X_K(t)$ and $\rho_h^{(K)} = R_K(h)/R_K(0)$.

Then we can show the following results (see Huzii [6]):

$$|\rho_h^{(K)}| \leq C(2\alpha)^h ,$$

$$\sum_{h=1}^{\infty} |\rho_h^{(K)} - \rho_h| \leq C_0 K(2\alpha)^K ,$$

where C and C_0 are constants, independent of K .

Let samples be $X_K(1), X_K(2), \dots, X_K(N)$. We shall construct $\{\hat{a}_K^{(K)}\}$, $\hat{\sigma}_\xi^2(K)$ and $\hat{f}^{(K)}(\lambda)$ by using $X_K(t)$ instead of $X(t)$.

Using $X_K(t)$ instead of $X(t)$, we can show the same results as Theorems 5 and 6 in Berk's paper [3]. So we can obtain that the distribution function of $\sqrt{N/K} \left(\hat{f}^{(K)}(\lambda) - f_K(\lambda) \right)$ tends to the normal distribution function with mean 0 and variance $2f(\lambda)^2$ when $0 < \lambda < 1/2$ or $4f(\lambda)^2$ when $\lambda = 0$ or $1/2$, if $\sqrt{NK} \alpha^K$ tends to 0 when N tends to infinity.

Now, for $X_K(t)$, we have the following lemma.

Lemma 3. There exists a solution $\tilde{f}^{(K)}(\lambda)$ for $f_K(\lambda)$, obtained by solving the likelihood equations, such that

$$g_N^{(K)}(\lambda) = \hat{f}^{(K)}(\lambda) - \tilde{f}^{(K)}(\lambda)$$

converges in probability to zero as N tends to infinity.

Proof. Let us consider the joint estimation of $(\sigma_\xi^2, a_1, a_2, \dots, a_K)$.

Let $\psi(X_N^{(K)})$ and $\psi(X_K^{(K)})$ be the density functions of

$\tilde{X}_N^{(K)} = (X_K(1), X_K(2), \dots, X_K(N))'$ and $\tilde{X}_K^{(K)} = (X_K(1), X_K(2), \dots, X_K(K))'$, respectively. Then we have

$$\psi(\tilde{X}_N^{(K)}) = \psi(\tilde{X}_K^{(K)}) \frac{1}{(2\pi)^{\frac{N-K}{2}} (\sigma_\xi^2)^{\frac{N-K}{2}}} \times \exp \left[-\frac{1}{2\sigma_\xi^2} \sum_{t=K+1}^N \left(X_K(t) + a_1 X_K(t-1) + \dots + a_K X_K(t-K) \right)^2 \right]$$

To prove the assertion of this lemma, it is enough to show that

$$\frac{\partial \log \psi(\tilde{X}_K^{(K)})}{\partial \sigma_\xi^2}$$

and

$$\frac{\partial \log \psi(\tilde{X}_K^{(K)})}{\partial a_k} \quad \text{for } k = 1, 2, \dots, K,$$

do not affect the solution of the likelihood equations.

Let $Q_K^{(K)}$ be the $K \times K$ matrix whose (k, ℓ) element is $\rho_{k-\ell}^{(K)}$ and $q_{k,\ell}^{(K)-1}$ be the (k, ℓ) element of $Q_K^{(K)-1}$. Then we have

$$\frac{1}{N} \frac{\partial \log \psi(\tilde{X}_K^{(K)})}{\partial a_k} = -\frac{1}{2} \frac{\rho_k^{(K)}}{\sigma_\xi^2} \frac{1}{N} \sum_{t=1}^K \sum_{s=1}^K q_{t,s}^{(K)-1} X_K(t) X_K(s)$$

$$\begin{aligned}
& - \frac{1}{2} \frac{\left(\sum_{k=0}^K a_k \rho_k^{(K)} \right)}{\sigma_\xi^2} \frac{1}{N} \sum_{t=1}^K \sum_{s=1}^K \frac{\partial q_{t,s}^{(K)-1}}{\partial a_k} X_K(t) X_K(s) \\
& + \frac{1}{2} \frac{\rho_k^{(K)}}{\left(\sum_{j=0}^K a_j \rho_j^{(K)} \right)} - \frac{1}{2} \frac{1}{N} \frac{\partial}{\partial a_k} \log |Q_K^{(K)}| .
\end{aligned}$$

But we can show

$$V \left(\frac{1}{N} \sum_{t=1}^K \sum_{s=1}^K q_{t,s}^{(K)-1} X_K(t) X_K(s) \right) = O \left(\frac{K^2}{N^2} \right) .$$

Now we have, for $1 \leq j \leq K-1$,

$$\left| \frac{\partial q_{t,s}^{(K)-1}}{\partial \rho_j^{(K)}} \right| \leq C_5 K, \quad \left| \frac{\partial \rho_j^{(K)}}{\partial a_k} \right| \leq C_6,$$

where C_5 and C_6 are constants, independent of K . So we can obtain

$$\left| \frac{\partial q_{t,s}^{(K)-1}}{\partial a_k} \right| = \left| \sum_{j=1}^{K-1} \frac{\partial q_{t,s}^{(K)-1}}{\partial \rho_j^{(K)}} \frac{\partial \rho_j^{(K)}}{\partial a_k} \right| \leq C_7 K^2,$$

where C_7 is a constant, independent of t, s and K . Therefore we can get

$$V \left(\frac{1}{N} \sum_{t=1}^K \sum_{s=1}^K \frac{\partial q_{t,s}^{(K)-1}}{\partial a_k} X_K(t) X_K(s) \right) = o \left(\frac{K^6}{N^2} \right).$$

We know

$$\begin{aligned} & \frac{1}{2} \frac{1}{N} \frac{\partial}{\partial a_k} \log |Q_K^{(K)}| - \frac{1}{2} \frac{\rho_k^{(K)}}{\left(\sum_{j=0}^K a_j \rho_j^{(K)} \right)} \\ &= E \left(- \frac{1}{2} \frac{\rho_k^{(K)}}{\sigma_\xi^2} \frac{1}{N} \sum_{t=1}^K \sum_{s=1}^K q_{t,s}^{(K)-1} X_K(t) X_K(s) \right) \\ &+ E \left(- \frac{1}{2} \frac{\left(\sum_{k=0}^K a_k \rho_k^{(K)} \right)}{\sigma_\xi^2} \frac{1}{N} \sum_{t=1}^K \sum_{s=1}^K \frac{\partial q_{t,s}^{(K)-1}}{\partial a_k} X_K(t) X_K(s) \right). \end{aligned}$$

So we obtain

$$\left| \frac{1}{2} \frac{1}{N} \frac{\partial}{\partial a_k} \log |Q_K^{(K)}| \right| = o \left(\frac{K^3}{N} \right).$$

Combining the above results, we can get

$$V \left(\frac{1}{N} \frac{\partial \log \psi(X_K^{(K)})}{\partial a_k} \right) = o \left(\frac{K^6}{N^2} \right).$$

And we also have

$$V \left(\frac{1}{N} \frac{\partial \log \psi(X_K^{(K)})}{\partial \sigma_\xi^2} \right) = O \left(\frac{K^2}{N^2} \right) .$$

Let $\hat{\tilde{R}}^{(K)}$ be the matrix which is constructed by $X_K(t)$ instead of $X(t)$ in $\hat{\tilde{R}}$, and let us denote

$$\hat{\tilde{\eta}}^{(K)} = \begin{pmatrix} \hat{\tilde{\eta}}_1^{(K)} \\ \hat{\tilde{\eta}}_2^{(K)} \\ \vdots \\ \hat{\tilde{\eta}}_K^{(K)} \end{pmatrix} = \hat{\tilde{R}}^{(K)-1} \tilde{\ell} ,$$

where

$$\tilde{\ell} = (\ell_1, \ell_2, \dots, \ell_K)'$$

$$= \left(\frac{1}{N} \frac{\partial \log \psi(X_K^{(K)})}{\partial a_1}, \frac{1}{N} \frac{\partial \log \psi(X_K^{(K)})}{\partial a_2}, \dots, \frac{1}{N} \frac{\partial \log \psi(X_K^{(K)})}{\partial a_K} \right) .$$

Then we have

$$\sqrt{K} \|\hat{\tilde{\eta}}^{(K)}\| \leq \sqrt{K} \|\hat{\tilde{R}}^{(K)-1} - \tilde{R}^{-1}\| \|\tilde{\ell}\| + \sqrt{K} \|\tilde{R}^{-1}\| \|\tilde{\ell}\| ,$$

where

$$\|\underline{\tilde{Y}}\|^2 = \sum_{k=1}^K y_k^2, \quad \underline{\tilde{Y}} = (y_1, y_2, \dots, y_K)',$$

and, for a $K \times K$ matrix $\underline{\tilde{B}}$, $\|\underline{\tilde{B}}\| = \sup \|\underline{\tilde{B}}\underline{\tilde{Y}}\|$ for $\|\underline{\tilde{Y}}\| \leq 1$. We can show in the same way as that of Berk [3] that $\|\underline{\tilde{R}}^{-1}\|$ is bounded and $\sqrt{K} \|\hat{\underline{\tilde{R}}}^{(K)-1} - \underline{\tilde{R}}^{-1}\|$ converges in probability to zero as N tends to infinity (see Berk[3]). Also we can show $\sqrt{K} \|\underline{\tilde{\lambda}}\|$ converges in probability to zero if K^4/N tends to zero when N tends to infinity. Therefore, under Assumption 2, $\sqrt{K} \|\hat{\underline{\tilde{\eta}}}^{(K)}\|$ converges in probability to zero. As we have

$$\left| \sum_{k=1}^K \hat{\eta}_k^{(K)} \cos 2\pi k \lambda \right| \leq \sqrt{K} \|\hat{\underline{\tilde{\eta}}}^{(K)}\|,$$

$$\left| \sum_{k=1}^K \hat{\eta}_k^{(K)} \sin 2\pi k \lambda \right| \leq \sqrt{K} \|\hat{\underline{\tilde{\eta}}}^{(K)}\|,$$

$\sum_{k=1}^K \hat{\eta}_k^{(K)} \cos 2\pi k \lambda$ and $\sum_{k=1}^K \hat{\eta}_k^{(K)} \sin 2\pi k \lambda$ converge in probability to zero under Assumption 2. Furthermore,

$$\begin{aligned} \left| \sum_{k=1}^K \hat{\eta}_k^{(K)} \hat{R}(0, k) \right| &\leq \left| \sum_{k=1}^K \hat{\eta}_k^{(K)} (\hat{R}(0, k) - R_K(k)) \right| + \left| \sum_{k=1}^K \hat{\eta}_k^{(K)} R_K(k) \right| \\ &\leq \|\hat{\underline{\tilde{\eta}}}^{(K)}\| \sqrt{\sum_{k=1}^K (\hat{R}(0, k) - R_K(k))^2} + \|\hat{\underline{\tilde{\eta}}}^{(K)}\| CR(0)/(1-2\alpha). \end{aligned}$$

Under Assumption 2, we can show $\|\hat{\eta}^{(K)}\|$ and $\sum_{k=1}^K \left(\hat{R}(0,k) - R(k) \right)^2$ converge in probability to zero. So under Assumption 2,

$\left| \sum_{k=1}^K \hat{\eta}_k^{(K)} \hat{R}(0,k) \right|$ converges in probability to zero.

From $\partial \log \psi(\tilde{X}_N^{(K)}) / \partial a_k = 0$ for $k = 1, 2, \dots, K$, we have

$$\sum_{j=1}^K a_j \hat{R}(0,j) = -\hat{R}(0,k) - 2\sigma_\xi^2 \ell_k$$

for $k = 1, 2, \dots, K$, and from $\partial \log \psi(\tilde{X}_N^{(K)}) / \partial \sigma_\xi^2 = 0$, we have

$$\sigma_\xi^2 = \hat{R}(0,0) + \sum_{k=1}^K a_k \hat{R}(0,k) + \frac{2\sigma_\xi^4}{N-K} \frac{\partial \log \psi(\tilde{X}_N^{(K)})}{\partial \sigma_\xi^2}.$$

Using the above results, we can obtain the assertion of this lemma.

In the following, let us put $\theta_0 = \sigma_\xi^2$, $\theta_1 = a_1$, $\theta_2 = a_2$, \dots , $\theta_K = a_K$ and $\hat{\theta}_0^{(K)} = \hat{\sigma}_\xi^2(K)$, $\hat{\theta}_1^{(K)} = \hat{a}_1^{(K)}$, $\hat{\theta}_2^{(K)} = \hat{a}_2^{(K)}$, \dots , $\hat{\theta}_K^{(K)} = \hat{a}_K^{(K)}$, for simplicity. We shall denote

$$\tilde{\theta}^{(K)} = (\theta_0, \theta_1, \dots, \theta_K)',$$

and

$$\hat{\theta}^{(K)} = \left(\hat{\theta}_0^{(K)}, \hat{\theta}_1^{(K)}, \dots, \hat{\theta}_K^{(K)} \right)' .$$

Let $\tilde{U}^{(K)}(X)$ and $\tilde{U}^{(K)}(X_K)$ be the $(K+1) \times (K+1)$ matrices whose

(i,j) elements are $E_X \left(\hat{\theta}_{i-1}^{(K)} - \theta_{i-1} \right) \left(\hat{\theta}_{j-1}^{(K)} - \theta_{j-1} \right)$ and

$E_{X_K} \left(\hat{\theta}_{i-1}^{(K)} - \theta_{i-1} \right) \left(\hat{\theta}_{j-1}^{(K)} - \theta_{j-1} \right)$, respectively, where, for example,

$E_X(\gamma)$ means the mean of the statistic γ for the process $X(t)$.

Let $\tilde{W}^{(K)}(X)^{-1}$ and $\tilde{W}^{(K)}(X_K)^{-1}$ be the $(K+1) \times (K+1)$ matrices whose (i,j) elements are

$$E_X \left(\frac{\partial \log \psi(X_N)}{\partial \theta_{i-1}} \frac{\partial \log \psi(X_N)}{\partial \theta_{j-1}} \right)$$

and

$$E_{X_K} \left(\frac{\partial \log \psi(X_N^{(K)})}{\partial \theta_{i-1}} \frac{\partial \log \psi(X_N^{(K)})}{\partial \theta_{j-1}} \right),$$

respectively. Let us denote the (i,j) element of $\tilde{W}^{(K)}(X)$ (or $\tilde{W}^{(K)}(X_K)$) as $w_{ij}^{(K)}(X)$ (or $w_{ij}^{(K)}(X_K)$).

Then we have the following lemma.

Lemma 4. It holds

$$\lim_{N \rightarrow \infty} \frac{y_K' \tilde{U}^{(K)}(X_K) y_K}{y_K' \tilde{W}^{(K)}(X_K) y_K} = 1.$$

for any sequence $\{\underline{y}_K\}$ of real vectors $\underline{y}_K = (y_1^{(K)}, y_2^{(K)}, \dots, y_K^{(K)})'$ such that $\underline{y}_K \neq (0, 0, \dots, 0)'$.

Proof. This result can be shown by the same method as the case when K is fixed.

$$\left. \frac{\partial \log \psi(\underline{X}_N^{(K)})}{\partial \theta_i} \right|_{\hat{\theta}} \approx 0$$

means

$$\begin{aligned} \left. \frac{1}{N} \frac{\partial \log \psi(\underline{X}_N^{(K)})}{\partial \theta_i} \right|_{\hat{\theta}} + \frac{1}{N} \sum_{j=0}^K (\hat{\theta}_j - \theta_j) \left\{ \left. \frac{\partial^2 \log \psi(\underline{X}_N^{(K)})}{\partial \theta_j \partial \theta_i} \right|_{\hat{\theta}} \right. \\ \left. + \frac{1}{2} \frac{1}{N} \sum_{\ell=0}^K (\hat{\theta}_\ell - \theta_\ell) \left. \frac{\partial^3 \log \psi(\underline{X}_N^{(K)})}{\partial \theta_\ell \partial \theta_j \partial \theta_i} \right|_{\hat{\theta} + \mu(\hat{\theta} - \theta)} \right\} = 0, \end{aligned}$$

where $0 < \mu < 1$. Let $\hat{\underline{B}}^{(K)}$ be the $(K+1) \times (K+1)$ matrix whose $(i+1, j+1)$ element $\hat{b}_{i+1, j+1}^{(K)}$ is

$$\hat{b}_{i+1, j+1}^{(K)} = \left. \frac{1}{N} \frac{\partial^2 \log \psi(\underline{X}_N^{(K)})}{\partial \theta_j \partial \theta_i} \right|_{\hat{\theta}} + \frac{1}{2} \frac{1}{N} \sum_{\ell=0}^K (\hat{\theta}_\ell - \theta_\ell) \left. \frac{\partial^3 \log \psi(\underline{X}_N^{(K)})}{\partial \theta_\ell \partial \theta_j \partial \theta_i} \right|_{\hat{\theta} + \mu(\hat{\theta} - \theta)}$$

and $\underline{B}^{(K)}$ be the $(K+1) \times (K+1)$ matrix whose $(i+1, j+1)$ element

$b_{i+1,j+1}$ is

$$b_{i+1,j+1} = \begin{cases} \frac{1}{2(\sigma_{\xi}^2)^2} & i=0, j=0 \\ 0 & i=0, j \geq 1 \text{ or } i \geq 1, j=0, \\ R(i-j) & i \geq 1, j \geq 1. \end{cases}$$

Then we can show, by the same method as that of Berk [3], that

$\sqrt{K} \|\hat{B}^{(K)-1} - \tilde{B}^{(K)-1}\|$ converges in probability to zero under

Assumption 2. Let us put

$$\ell_N^{(K)} = \left(\frac{1}{N} \frac{\partial \log \psi(X_N^{(K)})}{\partial \theta_0}, \frac{1}{N} \frac{\partial \log \psi(X_N^{(K)})}{\partial \theta_1}, \dots, \frac{1}{N} \frac{\partial \log \psi(X_N^{(K)})}{\partial \theta_K} \right)'.$$

Then we have

$$\hat{\theta}^{(K)} - \tilde{\theta}^{(K)} = - \hat{B}^{(K)-1} \ell_N^{(K)}.$$

And we can show

$$-\sqrt{\frac{N}{K}} \frac{1}{\|y_K\|} y_K' \hat{B}^{(K)-1} \ell_N^{(K)} - \left(-\sqrt{\frac{N}{K}} \frac{1}{\|y_K\|} y_K' \tilde{B}^{(K)-1} \ell_N^{(K)} \right)$$

converges in probability to zero when N tends to infinity. So

we can obtain

$$\lim_{N \rightarrow \infty} \frac{N}{K} \frac{1}{\|\underline{y}_K\|^2} \underline{y}_K' \underline{U}^{(K)}(\underline{x}_K) \underline{y}_K = \lim_{N \rightarrow \infty} \frac{N}{K} \frac{1}{\|\underline{y}_K\|^2} \underline{y}_K' \underline{B}^{(K)-1} \left(E \underline{\ell}_N^{(K)} \underline{\ell}_N^{(K)'} \right) \underline{B}^{(K)-1} \underline{y}_K.$$

Therefore, we can get

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{\underline{y}_K' \underline{U}^{(K)}(\underline{x}_K) \underline{y}_K}{\underline{y}_K' \underline{W}^{(K)}(\underline{x}_K) \underline{y}_K} &= \lim_{N \rightarrow \infty} \frac{\frac{N}{K} \frac{1}{\|\underline{y}_K\|^2} \underline{y}_K' \underline{B}^{(K)-1} \left(E \underline{\ell}_N^{(K)} \underline{\ell}_N^{(K)'} \right) \underline{B}^{(K)-1} \underline{y}_K}{\frac{N}{K} \frac{1}{\|\underline{y}_K\|^2} \underline{y}_K' \underline{W}^{(K)}(\underline{x}_K) \underline{y}_K} \\ &= 1 \end{aligned}$$

by using Assumption 2.

Now we shall consider the joint estimation of $\underline{\theta}^{(K)}$, i.e., $(\sigma_\xi^2, a_1, a_2, \dots, a_K)$, by \underline{x}_N . Let $\underline{\tilde{\theta}}^{(K)} = (\tilde{\theta}_0^{(K)}, \tilde{\theta}_1^{(K)}, \dots, \tilde{\theta}_K^{(K)})$ be arbitrary joint estimators of $\underline{\theta}^{(K)}$ which have the following properties.

For any sequence $\{\underline{y}_K\}$ of real vectors $\underline{y}_K = (y_1^{(K)}, y_2^{(K)}, \dots, y_K^{(K)})'$ such that $\underline{y}_K \neq (0, 0, \dots, 0)'$,

$$(6) \quad \lim_{N \rightarrow \infty} \frac{\underline{y}_K' \underline{\tilde{B}}^{(K)} \underline{\tilde{U}}^{(K)}(\underline{x})^{-1} \underline{y}_K}{\underline{y}_K' \underline{\tilde{U}}^{(K)}(\underline{x})^{-1} \underline{y}_K} = 0$$

and

$$(7) \quad \lim_{N \rightarrow \infty} \frac{\tilde{y}_K' \tilde{B}^{(K)} \tilde{U}^{(K)}(X)^{-1} \tilde{B}^{(K)} \tilde{y}_K}{\tilde{y}_K' \tilde{U}^{(K)}(X)^{-1} \tilde{y}_K} = 0 ,$$

where $\tilde{B}^{(K)}$ is the $(K+1) \times (K+1)$ matrix whose (k, ℓ) element is $\partial b_{k-1}^{(K)} / \partial \theta_{\ell-1}$, $b_{k-1}^{(K)} = E(\tilde{\theta}_{k-1}^{(K)} - \theta_{k-1})$, and $\tilde{U}^{(K)}(X)$ is the $(K+1) \times (K+1)$ matrix, the (i, j) element of which is $E_X(\tilde{\theta}_{i-1}^{(K)} - \theta_{i-1})(\tilde{\theta}_{j-1}^{(K)} - \theta_{j-1})$. In this case, we have (see Cramér [4])

$$\begin{aligned} \tilde{y}_K' \tilde{W}^{(K)}(X)^{-1} \tilde{y}_K &\geq \tilde{y}_K' \tilde{U}^{(K)}(X)^{-1} \tilde{y}_K \\ &+ \tilde{y}_K' \left(\tilde{B}^{(K)} \tilde{U}^{(K)}(X)^{-1} + \tilde{U}^{(K)}(X)^{-1} \tilde{B}^{(K)} + \tilde{B}^{(K)} \tilde{U}^{(K)}(X)^{-1} \tilde{B}^{(K)} \right) \tilde{y}_K . \end{aligned}$$

So we can obtain

$$\lim_{N \rightarrow \infty} \frac{\tilde{y}_K' \tilde{W}^{(K)}(X)^{-1} \tilde{y}_K}{\tilde{y}_K' \tilde{U}^{(K)}(X)^{-1} \tilde{y}_K} \geq 1 .$$

This means

$$\lim_{N \rightarrow \infty} \frac{\tilde{y}_K' \tilde{U}^{(K)}(X) \tilde{y}_K}{\tilde{y}_K' \tilde{W}^{(K)}(X) \tilde{y}_K} \geq 1 .$$

For our $\hat{\theta}^{(K)}$, we can show the conditions (6) and (7) are satisfied under Assumption 2. In fact, $b_{k-1}^{(K)}$ is an analytic function of $\tilde{\theta} = (\theta_0, \theta_1, \theta_2, \dots)'$, and we know that $b_{k-1}^{(K)}$ tends to zero uniformly in $1 \leq k \leq K+1$ when N tends to infinity. If we put $\theta_{K+1} = \theta_{K+2} = \dots = 0$, which is the case of an autoregressive process of order K , we can show $\partial b_{k-1}^{(K)} / \partial \theta_{\ell-1}$ converges to zero as N tends to infinity, and this convergence does not depend on the values of $\theta_0, \theta_1, \dots$. Therefore, if we express $b_{k-1}^{(K)}$ as $b_{k-1}^{(K)} = \tilde{b}_{k-1}^{(K)} + \tilde{\tilde{b}}_{k-1}^{(K)}$, where $\tilde{\tilde{b}}_{k-1}^{(K)} = o(1/N^\delta)$ ($\delta > 0$), we can consider $\tilde{b}_{k-1}^{(K)}$ is an analytic function of $\theta_{K+1}, \theta_{K+2}, \dots$. So if

$$\lim_{N \rightarrow \infty} b_{k-1}^{(K)} = 0,$$

then

$$\lim_{N \rightarrow \infty} \frac{\partial b_{k-1}^{(K)}}{\partial \theta_{\ell-1}} = 0$$

and this convergence does not depend on k and ℓ , because $\{a_k\}$ satisfies Assumption 1 and $b_{k-1}^{(K)}$ tends to zero uniformly in $1 \leq k \leq K+1$. We have

$$(K+1)^2 \max_{1 \leq k, \ell \leq K+1} \left| \left(\partial b_{k-1}^{(K)} / \partial \theta_{\ell-1} \right) \right|$$

tends to zero as N tends to infinity. Therefore, we can see

$$\lim_{N \rightarrow \infty} \frac{\tilde{y}_K' \tilde{B}^{(K)} \tilde{U}^{(K)}(X)^{-1} \tilde{y}_K}{\tilde{y}_K' \tilde{U}^{(K)}(X)^{-1} \tilde{y}_K} = 0, \quad \lim_{N \rightarrow \infty} \frac{\tilde{y}_K' \tilde{B}^{(K)} \tilde{U}^{(K)}(X)^{-1} \tilde{B}^{(K)} \tilde{y}_K}{\tilde{y}_K' \tilde{U}^{(K)}(X)^{-1} \tilde{y}_K} = 0.$$

In the following, we shall consider the difference between $\tilde{W}^{(K)}(X)$ and $\tilde{W}^{(K)}(X_K)$. This means we have to evaluate

$$\tilde{y}_K' \tilde{W}^{(K)}(X) \tilde{y}_K - \tilde{y}_K' \tilde{W}^{(K)}(X_K) \tilde{y}_K$$

for any sequence $\{\tilde{y}_K\}$ of real vectors $\tilde{y}_K = (y_1^{(K)}, y_2^{(K)}, \dots, y_K^{(K)})'$. But for this purpose, we shall compare $\tilde{W}^{(K)}(X)^{-1}$ with $\tilde{W}^{(K)}(X_K)^{-1}$.

In the first place, we shall compare $E_X \left(\partial \log \psi(\tilde{X}_N) / \partial \theta_{i-1} \right) \left(\partial \log \psi(\tilde{X}_N) / \partial \theta_{j-1} \right)$ with $E_{X_K} \left(\partial \log \psi(\tilde{X}_N^{(K)}) / \partial \theta_{i-1} \right) \left(\partial \log \psi(\tilde{X}_N^{(K)}) / \partial \theta_{j-1} \right)$.

Let us put $\tilde{\theta} = (\theta_0, \theta_1, \theta_2, \dots)' = (\sigma_\xi^2, a_1, a_2, \dots)'$ and

$F_{ij}(\tilde{\theta}) = E_X \left(\partial \log \psi(\tilde{X}_N) / \partial \theta_{i-1} \right) \left(\partial \log \psi(\tilde{X}_N) / \partial \theta_{j-1} \right)$. Then we have

$$F_{ij}(\tilde{\theta}) \Big|_{\tilde{\theta} = \tilde{\theta}_K} = E_{X_K} \left(\partial \log \psi(\tilde{X}_N^{(K)}) / \partial \theta_{i-1} \right) \left(\partial \log \psi(\tilde{X}_N^{(K)}) / \partial \theta_{j-1} \right),$$

where $\tilde{\theta}_K = (\sigma_\xi^2, a_1, a_2, \dots, a_K, 0, 0, \dots)$, and

$$F_{ij}(\theta) = F_{ij}(\tilde{\theta}) \Big|_{\tilde{\theta} = \theta_K} + \sum_{k=K+1}^{\infty} a_k \frac{\partial F_{ij}(\tilde{\theta})}{\partial \theta_k} \Big|_{\theta_K + \tau(\theta - \theta_K)},$$

where $0 < \tau < 1$. By differentiating both sides of the equations

$$\sum_{j=1}^N q_{ij}^{-1} \rho_{i'-j} = \begin{cases} 1 & i = i' \\ 0 & i \neq i' \end{cases}$$

with respect to ρ_h , we can obtain

$$\left| \frac{\partial q_{ij}^{-1}}{\partial \rho_h} \right| \leq C_8 N, \quad \left| \frac{\partial^2 q_{ij}^{-1}}{\partial \rho_h \partial \rho_{h'}} \right| \leq C_8' N^2, \quad 1 \leq h, h' \leq N-1$$

where C_8 and C_8' are constants, independent of i, j, h, h' and N . And we have

$$\sum_{h=1}^{\infty} \left| \frac{\partial \rho_h}{\partial a_k} \right| \leq C_9, \quad \sum_{h=1}^{\infty} \left| \frac{\partial^2 \rho_h}{\partial a_{K+l} \partial a_k} \right| < C_9',$$

where C_9 and C_9' are constants, independent of l, h, k and K .

So we can get

$$\begin{aligned} \left| \frac{\partial q_{ij}^{-1}}{\partial a_{K+l}} \right| &= \left| \sum_{h=1}^{N-1} \frac{\partial q_{ij}^{-1}}{\partial \rho_h} \frac{\partial \rho_h}{\partial a_{K+l}} \right| \\ &\leq C_{10} N^2, \quad \text{for } 1 \leq l, \end{aligned}$$

where C_{10} is a constant, independent of i, j, K and ℓ .

And we also have

$$\begin{aligned}
 \left| \frac{\delta^2 q_{ij}^{-1}}{\partial a_{K+\ell} \partial a_m} \right| &= \left| \sum_{h=1}^{N-1} \left(\frac{\delta^2 q_{ij}^{-1}}{\partial a_{K+\ell} \partial \rho_h} \frac{\partial \rho_h}{\partial a_m} + \frac{\partial q_{ij}^{-1}}{\partial \rho_h} \frac{\delta^2 \rho_h}{\partial a_{K+\ell} \partial a_m} \right) \right| \\
 &= \left| \sum_{h=1}^{N-1} \left\{ \left(\sum_{h'=1}^{N-1} \frac{\delta^2 q_{ij}^{-1}}{\partial \rho_h \partial \rho_{h'}} \frac{\partial \rho_{h'}}{\partial a_{K+\ell}} \right) \frac{\partial \rho_h}{\partial a_m} + \frac{\partial q_{ij}^{-1}}{\partial \rho_h} \frac{\delta^2 \rho_h}{\partial a_{K+\ell} \partial a_m} \right\} \right| \\
 &\leq C_{11} N^2,
 \end{aligned}$$

where C_{11} is a constant independent of m, ℓ, i, j and N .

Using the above results, we can get

$$\left| \sum_{k=K+1}^{\infty} a_k \frac{\partial F_{ij}(\theta)}{\partial \theta_k} \right|_{\theta_K + \tau(\theta - \theta_K)} \leq C_{12} N^9 \alpha^K,$$

where C_{12} is a constant, independent of i, j , and N .

So we can obtain

$$\left\| \tilde{W}^{(K)}(X)^{-1} - \tilde{W}^{(K)}(X_K)^{-1} \right\| \leq \sqrt{C_{12} N^9 K^2 \alpha^K},$$

the right hand side of which tends to zero as N tends to infinity.

Now we have

$$\begin{aligned}
\|\tilde{W}^{(K)}(X) - \tilde{W}^{(K)}(X_K)\| &= \|\tilde{W}^{(K)}(X) \left(\tilde{W}^{(K)}(X_K)^{-1} - \tilde{W}^{(K)}(X)^{-1} \right) \tilde{W}^{(K)}(X_K)\| \\
&\leq \|\tilde{W}^{(K)}(X)\| \|\tilde{W}^{(K)}(X_K)^{-1} - \tilde{W}^{(K)}(X)^{-1}\| \|\tilde{W}^{(K)}(X_K)\|.
\end{aligned}$$

We know

$$\tilde{W}^{(K)}(X_K) = \frac{\sigma_\xi^2}{N-K} \begin{pmatrix} 2\sigma_\xi^2 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \tilde{R}^{-1} & \\ 0 & & & \end{pmatrix}.$$

So we can get

$$\|\tilde{W}^{(K)}(X_K)\| \leq \sqrt{\frac{C_{13} K^3}{(N-K)^2}},$$

where C_{13} is a constant, independent of N . Using the above facts, we can obtain

$$\|\tilde{W}^{(K)}(X) - \tilde{W}^{(K)}(X_K)\| \leq \frac{\|\tilde{W}^{(K)}(X_K)\|^2 \|\tilde{W}^{(K)}(X_K)^{-1} - \tilde{W}^{(K)}(X)^{-1}\|}{1 - \|\tilde{W}^{(K)}(X_K)^{-1} - \tilde{W}^{(K)}(X)^{-1}\| \|\tilde{W}^{(K)}(X_K)\|},$$

the right hand side of which converges to zero as N tends to infinity.

Therefore we have

$$\lim_{N \rightarrow \infty} \left| \left(\tilde{y}_K / \|\tilde{y}_K\| \right)' \left(\tilde{W}^{(K)}(X) - \tilde{W}^{(K)}(X_K) \right) \left(\tilde{y}_K / \|\tilde{y}_K\| \right) \right|$$

$$\leq \lim_{N \rightarrow \infty} \| \tilde{W}^{(K)}(X) - \tilde{W}^{(K)}(X_K) \| = 0 .$$

Consequently, we can obtain

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{\tilde{y}_K' \tilde{W}^{(K)}(X_K) \tilde{y}_K}{\tilde{y}_K' \tilde{W}^{(K)}(X) \tilde{y}_K} &= \lim_{N \rightarrow \infty} \frac{\left(\tilde{y}_K / \|\tilde{y}_K\| \right)' \tilde{W}^{(K)}(X_K) \left(\tilde{y}_K / \|\tilde{y}_K\| \right)}{\left(\tilde{y}_K / \|\tilde{y}_K\| \right)' \tilde{W}^{(K)}(X) \left(\tilde{y}_K / \|\tilde{y}_K\| \right)} \\ &= \lim_{N \rightarrow \infty} \frac{\tilde{z}_K' \tilde{W}^{(K)}(X_K) \tilde{z}_K}{\tilde{z}_K' \tilde{W}^{(K)}(X) \tilde{z}_K} \\ &= \lim_{N \rightarrow \infty} \frac{\tilde{z}_K' \tilde{W}^{(K)}(X) \tilde{z}_K + \tilde{z}_K' \left(\tilde{W}^{(K)}(X_K) - \tilde{W}^{(K)}(X) \right) \tilde{z}_K}{\tilde{z}_K' \tilde{W}^{(K)}(X) \tilde{z}_K} \\ &= 1 , \end{aligned}$$

where $\tilde{z}_K = \tilde{y}_K / \|\tilde{y}_K\|$.

Summarizing the above results, we have the following lemma.

Lemma 5. Let $\{\tilde{y}_K\}$ be any sequence of real vectors

$\tilde{y}_K = \left(y_1^{(K)}, y_2^{(K)}, \dots, y_K^{(K)} \right)'$ such that $\tilde{y}_K \neq (0, 0, \dots, 0)'$.

Then

$$\lim_{N \rightarrow \infty} \frac{\tilde{y}_K' \tilde{W}^{(K)}(X_K) \tilde{y}_K}{\tilde{y}_K' \tilde{W}^{(K)}(X) \tilde{y}_K} = 1 .$$

Now we have

$$\begin{aligned} \hat{f}^{(K)}(\lambda) - f_K(\lambda) &= \frac{\partial f_K(\lambda)}{\partial \sigma_\xi^2} \left(\hat{\sigma}_\xi^2(K) - \sigma_\xi^2 \right) + \sum_{k=1}^K \frac{\partial f_K(\lambda)}{\partial a_k} \left(\hat{a}_k^{(K)} - a_k \right) \\ &+ O \left\{ \left[\left(\hat{\sigma}_\xi^2(K) - \sigma_\xi^2 \right)^2 + \sum_{k=1}^K \left(\hat{a}_k^{(K)} - a_k \right)^2 \right]^{\frac{1}{2}} \right\}, \end{aligned}$$

where the last term converges in probability to zero more rapidly than the first two terms. So we can consider

$$\hat{f}^{(K)}(\lambda) - f_K(\lambda) \sim \frac{\partial f(\lambda)}{\partial \sigma_\xi^2} \left(\hat{\sigma}_\xi^2(K) - \sigma_\xi^2 \right) + \sum_{k=1}^K \frac{\partial f(\lambda)}{\partial a_k} \left(\hat{a}_k^{(K)} - a_k \right).$$

Let us put $\tilde{f}^{(K)} = \left(\partial f(\lambda)/\partial \sigma_\xi^2, \partial f(\lambda)/\partial a_1, \dots, \partial f(\lambda)/\partial a_K \right)'$, and we shall take $\tilde{f}^{(K)}$ as \tilde{y}_K in Lemma 5. Then the above discussion means, for sufficiently large N ,

$$\begin{aligned} \frac{N}{K} E_X \left(\hat{f}^{(K)}(\lambda) - f(\lambda) \right)^2 &\approx \frac{N}{K} E_{X_K} \left(\hat{f}^{(K)}(\lambda) - f(\lambda) \right)^2 \\ &\approx \frac{N}{K} \tilde{f}^{(K)'} \tilde{W}^{(K)}(X_K) \tilde{f}^{(K)} \\ &\approx \frac{N}{K} \tilde{f}^{(K)'} \tilde{W}^{(K)}(X) \tilde{f}^{(K)}. \end{aligned}$$

So, when N is sufficiently large, we have

$$\frac{N}{K} E_X \left(\hat{f}^{(K)}(\lambda) - f(\lambda) \right)^2 \approx \frac{N}{K} \tilde{f}^{(K)'} \tilde{W}^{(K)}(X) \tilde{f}^{(K)},$$

which is asymptotically minimum in a class of estimates $\tilde{f}^{(K)}(\lambda)$ of $f_K(\lambda)$. Every element of this class can be expressed as

$$\begin{aligned} \tilde{f}^{(K)}(\lambda) = & f(\lambda) + \frac{\partial f}{\partial \sigma_\xi^2} \left(\tilde{\sigma}_\xi^2 - \sigma_\xi^2 \right) + \sum_{k=1}^K \frac{\partial f}{\partial a_k} \left(\tilde{a}_k^{(K)} - a_k \right) \\ & + O \left(\sqrt{\left(\tilde{\sigma}_\xi^2 - \sigma_\xi^2 \right)^2 + \sum_{k=1}^K \left(\tilde{a}_k^{(K)} - a_k \right)^2} \right), \end{aligned}$$

where $\tilde{\sigma}_\xi^2, \tilde{a}_1^{(K)}, \dots, \tilde{a}_K^{(K)}$ are arbitrary estimates of $\sigma_\xi^2, a_1, \dots, a_K$, which satisfy conditions (6) and (7).

Summarizing the above results, we have the following theorem.

Theorem 2. Let $X(t)$ be a stationary normal process with mean zero and satisfy Assumption 1, and let N and K satisfy Assumption 2. Then we have

$$\lim_{N \rightarrow \infty} \frac{\frac{N}{K} E \left(\hat{f}^{(K)}(\lambda) - f(\lambda) \right)^2}{\frac{N}{K} \tilde{f}^{(K)'} \tilde{W}^{(K)}(X) \tilde{f}^{(K)}} = 1.$$

In general, there would be many ways to define the joint efficiency of estimators of infinite dimensional unknown parameters. In the above discussion, we have defined the efficiency as the limit

of the sequence of efficiencies of joint estimators of $(K+1)$ dimensional unknown parameters $(\sigma_{\xi}^2, a_1, a_2, \dots, a_K)$ for $K = 1, 2, 3, \dots$. We have shown $\hat{f}^{(K)}(\lambda)$ has the efficiency in this sense for the case we have treated.

Acknowledgement

The author wishes to thank Professors Theodore W. Anderson and David R. Brillinger for their many helpful comments.

REFERENCES

- [1] Akaike, H. (1969), "Power spectrum estimation through autoregressive model fitting". Ann. Inst. Statist. Math., 21, 407-419.
- [2] Anderson, T. W. (1971), The Statistical Analysis of Time Series, John Wiley & Sons, Inc., New York.
- [3] Berk, K. N. (1974), "Consistent autoregressive spectral estimates". The Annals of Statistics, 2, 489-502.
- [4] Cramér, H. (1946), "A contribution to the theory of statistical estimation", Skand. Aktuarietidskr., 29, 85-94.
- [5] Gersch, W. and Sharpe, D. R. (1973), "Estimation of power spectra with finite-order autoregressive models. IEEE Trans. Automatic Control. AC-18, 367-369.
- [6] Huzii, M., "On an efficient estimate of correlogram". Essays in Probability and Statistics, Tokyo. To appear.
- [7] Jones, R. H. (1974), "Identification and autoregressive spectrum estimation". IEEE Trans. Automatic Control. AC-19, 894-898.
- [8] Parzen, E. (1969), "Multiple time series modeling", Multivariate Analysis II (ed. by P. R. Krishnaiah), Academic Press, New York, 389-409.
- [9] Wise, J. (1955), "The autocorrelation function and the spectral density function". Biometrika 42, 6, 151-159.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.
2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.
3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.
4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.
5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.
6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.
7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.
8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.
9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.
10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.
11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.
12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.
13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.
14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.
15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.
16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time Series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations,"
Steffen L. Lauritzen, May, 1974.
18. "On the Interrelationships among Sufficiency, Total Sufficiency and
Some Related Concepts," Steffen L. Lauritzen, June, 1974.
19. "Statistical Inference for Multiply Truncated Power Series Distributions,"
T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models
in the Time and Frequency Domains," T. W. Anderson, June 1975.
21. "Asymptotic Properties of Some Estimators in Moving Average Models,"
Raul Pedro Mentz, September 8, 1975.
22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting,"
Mituaki Huzii, February 1976.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 22	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON A SPECTRAL ESTIMATE OBTAINED BY AN AUTOREGRESSIVE MODEL FITTING		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) MITUAKI HUZII		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE February 1976
		13. NUMBER OF PAGES 37
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; Distribution Unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) efficiency, spectral estimate, autoregressive model fitting		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) SEE REVERSE SIDE		

DD FORM 1473
1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

We shall consider a stationary Gaussian process $X(t)$ which is expressed as an autoregressive process of infinite order. We shall try to fit an autoregressive model of finite order K for this process and obtain an estimate for the spectral density. Berk [3] has shown the consistency and the asymptotic normality of this estimate under some conditions. We shall show this estimate has an asymptotically efficient property in a sense under some conditions which are stronger than Berk's conditions. The main conditions are that the autoregressive coefficient a_k of $X(t-k)$ ($k \geq 1$) is not greater than α^k in absolute value, where $0 < \alpha < 1/2$, and that K tends to infinity and K^4/N , N being the sample size, and $N^9 K^2 \alpha^K$ tend to 0 when N tends to infinity.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)